34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Feature Accentuation in Phosphenated Images*

Filiz Isabell Kiral-Kornek*,[1,2], Craig O. Savage[1,2], David B. Grayden[1,2,3,4], and Anthony N. Burkitt[1,2,3,4]

Abstract— **We present and evaluate different approaches to feature accentuation in phosphenated images for different image resolutions. The goal of this study is to find methods to attract an implantee's visual attention to important image content like faces, obstacles or road signs. We do this by defining an important region in the image and accentuating it by either increasing the brightness of outlining phosphenes or by using elliptical phosphenes to circumscribe the feature. While we only see limited benefit of ellipse phosphenes for a high-resolution prosthesis, the use of elliptical phosphenes of different orientations is a promising way to highlight features in a low-resolution phosphene representation of an image.**

I. INTRODUCTION

Since it has been shown that electrical stimulation of retinal ganglion cells can lead to visual perceptions in patients, various groups have been working on developing retinal implants to restore vision for people with vision loss due to photoreceptor decay. The visual impression these patients perceive is often described as bright or dark dots called phosphenes with different brightness levels [1]. The brightness levels of individual phosphenes are known to depend upon stimulation parameters such as current amplitude, pulse width, and interphase gap [2], [3], [4]. Recent research, however, suggests that not only brightness but also the shape of phosphenes can be influenced. This can either be done by changing the stimulation order of neighboring electrodes [5] or by choosing different return electrodes [6], [7]. In this paper, we examine possibilities to use these results to accentuate features in static images. This approach may be used to improve feature detection such as faces, obstacles or road signs. As blind patients often feel isolated [8], a main goal for high-acuity implants is face detection. Thus we will demonstrate our idea for the task of counting faces in an image. How well this can be done using conventional stimulation strategies strongly depends on the resolution, i.e. the number of electrodes of an implant. While available retinal implants have 60 electrodes, electrode arrays with 1000 - 1500 electrodes are under development [9], [10]. Therefore, we compare three different stimulation strategies for two different electrode arrays with respect to their ability to accentuate previously defined features in a static image. In a first attempt to evaluate these strategies, we use a bottom-up saliency algorithm that has been developed and implemented by Walther [11].

Other possibilities to highlight specific features of an image include zooming into faces [12] or modulating the brightness of the groundplane [13]. While these strategies are limited to use only the grayscale information of the image, by using elliptical phosphenes we suggest another dimension. This brings the advantage that we are not changing the luminance information of the input image.

The remainder of this document is structured as follows: in Section II-A, we describe different methods that we use to create phosphenated images with different feature accentuation techniques. We briefly outline the saliency algorithm that we use to compare the different methods in Section II-B. Results are shown in Section III, which is followed by a discussion in Section IV in which we also outline future work. Section V concludes this document.

II. METHODS

A. Phosphenization

We simulate phosphene vision under the assumptions that current density is the determining factor for neural activity and that impedance is homogenous. Thus, each phosphene is modeled with a two-dimensional Gaussian. In the simplest case, each electrode will elicit exactly one round phosphene. By varying the return electrode allocation, we incorporate cross talk and can thus produce elliptical phosphenes [7] that are modeled as non-overlapping but elongated two-dimensional Gaussians. We use six bits that correspond to 64 current steps to define the brightness level of each phosphene.

In this paper, we compare three different stimulation strategies with respect to their ability to accentuate features, specifically faces. An example image is shown in Fig. 1(a) [14].

1) Conventional phosphenization: We compare two different methods to the conventional method of phosphenization. Round phosphenes are used to display the grayscale information of an image. To minimize the impact of noise, the input image is blurred with a two-dimensional Gaussian. The blurred image is subsampled so that the number of values matches the number of electrodes and each brightness value is than again convolved with a two-dimensional Gaussian. A stimulation strategy based on this perceptual model only uses brightness to convey information. Sample images are shown in Fig. 1(b) for 8×8 electrodes and Fig. 1(f) for 32×32 electrodes.

*This research was supported by the Australian Research Council (ARC) through its Special Research Initiative (SRI) in Bionic Vision Australia (BVA).

[1]NeuroEngineering Laboratory, Department of Electrical & Electronic Engineering, The University of Melbourne VIC, 3010, Australia

[2]Centre for Neural Engineering, The University of Melbourne

[3]NICTA, c/- Department of Electrical & Electronic Engineering, The University of Melbourne, Victoria, 3010, Australia

[4]Bionics Institute, 384-388 Albert St, East Melbourne VIC 3010, Australia

*Correspondence to be sent to:
i.kiral-kornek@student.unimelb.edu.au

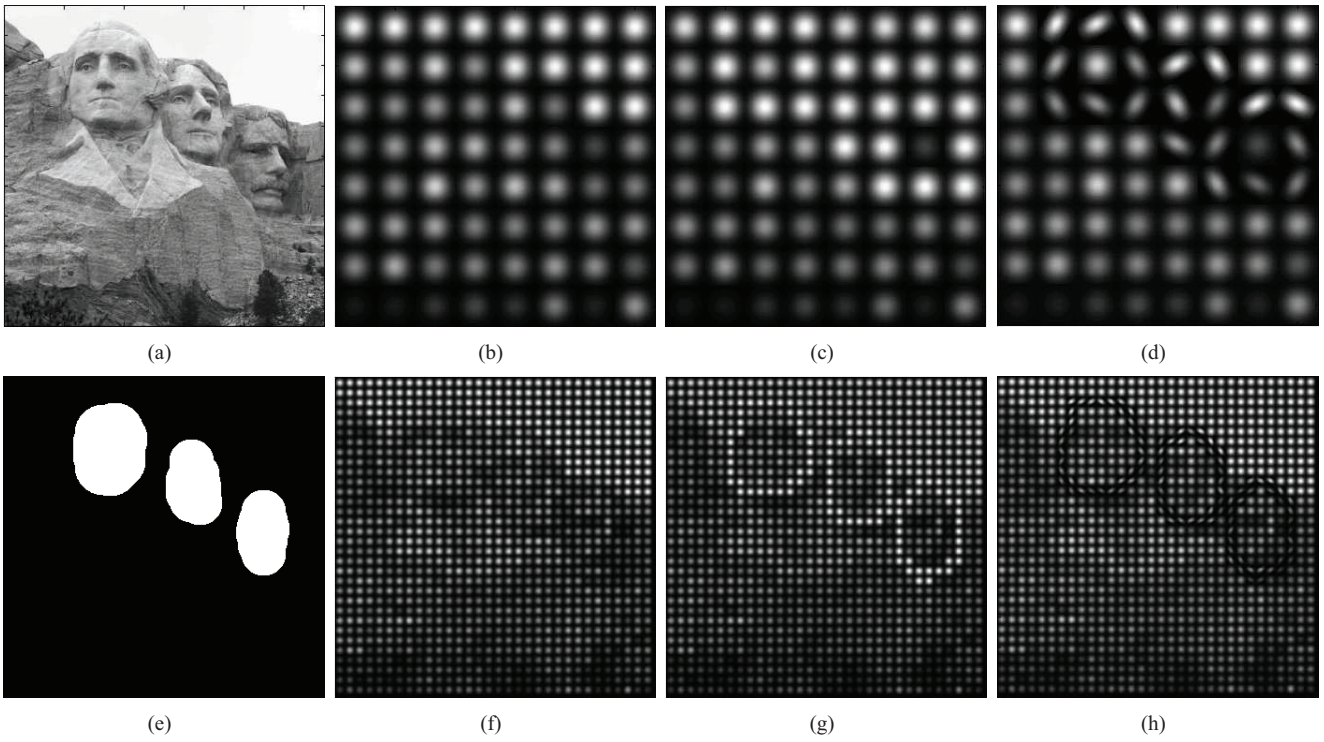

Fig. 1: Original grayscale image (a) [14], accentuation map (e), the output the conventional phosphenization method (b), bright outlining (c), elliptical outlining (d) for 8×8 electrodes and the output the conventional phosphenization method (f), bright outlining (g), elliptical outlining (h) for 32×32 electrodes.

2) Bright outline: The simplest way to accentuate a feature in an image is to make it brighter than the surroundings. This, however, creates two problems. First, we lose grayscale information of the accentuated part of the image, which we can minimize by just brightening the outline of the face. The second problem is that this strategy does not work when the feature we want to highlight is on a bright background. Making the outline darker for these cases could solve this issue, but would also result in ambiguous information for the user on what to focus on. To brighten the outline, we create a mask such as the one displayed in Fig. 1(e) for each input image to label the regions of interest. In a later stage, this can be done using image segmentation techniques. We detect the edge of the mask using a Sobel edge detection algorithm. It is then phosphenated as described in Section II-A with the exception that when an edge falls in the region of a certain phosphene, we display this phosphene at a maximally bright level. Phosphene images created with this method are depicted in Fig. 1(c) for 8×8 electrodes and Fig. 1(g) for 32×32 electrodes.

3) Elliptical outline: To minimize information loss due to brightness level adjustment, we make use of another dimension: the shape of single phosphenes. It has been shown that it is theoretically possible to deliberately create elliptical phosphenes [6], [7]. To highlight the outline of the region of interest with tilted ellipses, we need to find the direction of the edge first. We do this by filtering the contour of the mask (Fig. 1(e)) retrieved via edge detection with directed filters and label each image patch with its major direction. The corresponding phosphene is then created using an elongated 2-dimensional Gaussian. Sample images can be found in Fig. 1(d) and Fig. 1(h) for different electrode array sizes.

B. Saliency

Ultimately, we will evaluate the different methods by presenting the images to observers. For a first computational evaluation, we chose to compare the bottom-up saliency of highlighted areas. It should be noted that the algorithm used will not be trained to react to faces or any specific pattern including contour and closure effects. Nevertheless, it is biologically plausible and will give us an idea how well the highlighted area pops out. This, of course, can only be a first attempt of evaluation, but will tell us if highlighted areas are visually attractive. The algorithm used has first been described by Itti *et al.* [15] and has been implemented as a MATLAB toolbox [11]. Thus, we only give a brief description here.

In the first step, brightness, color, and orientation of an image are separately lowpass filtered at different scales to minimize the influence of noise. Across all scales, center-surround differences akin to those found in the visual pathway are calculated to obtain contrast sensitive feature maps. Feature maps are normalized iteratively using the algorithm of Itti and Koch [16]. The normalized feature maps are then combined across scale and finally added to one saliency map. Center and surround scales as well as the level at which the

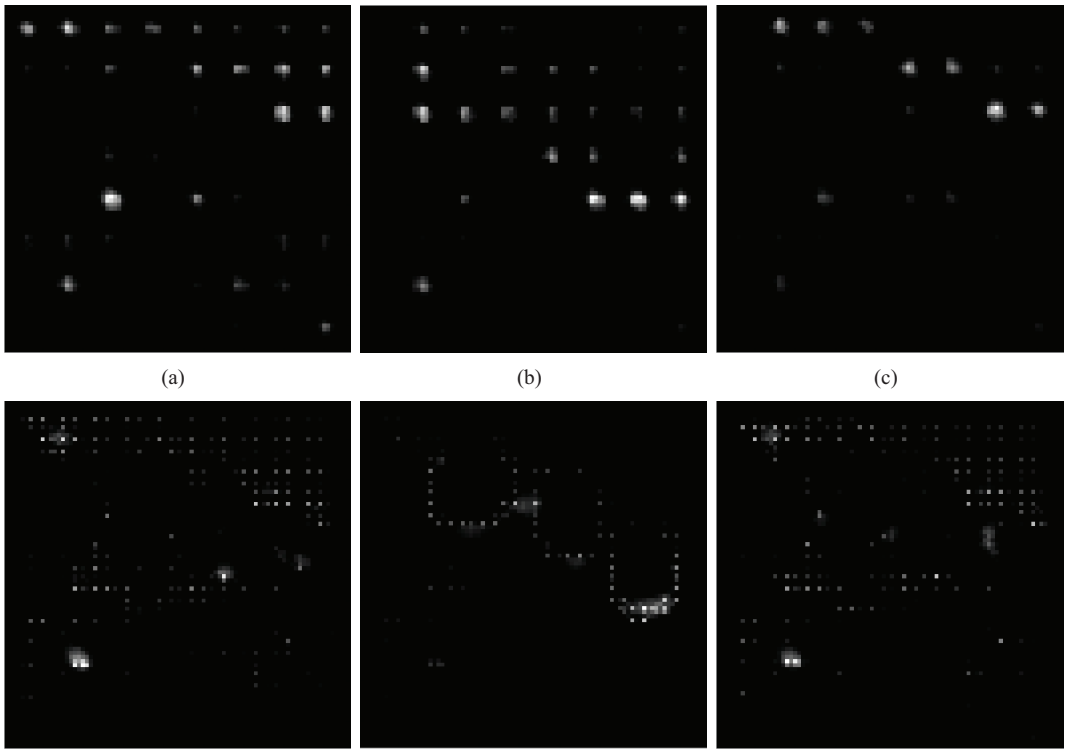

Fig. 2: Saliency map for images created with the conventional phosphenization method (a), bright outlining (b), elliptical outlining (c) for 8×8 electrodes saliency of images created with the conventional phosphenization method (d), bright outlining (e), elliptical outlining (f) for 32×32 electrodes.

saliency map is computed can be chosen freely. While the literature suggests that these levels do not matter too much for normal images [17], we already pointed out that while looking at a phosphenated image, the choice of the levels plays a bigger role [18]. Our input images have a resolution of 344×344 pixels. We set the lowest surround scale to 2 and the highest to 3. The center-surround difference is between 1 and 2 and saliency is computed at scale 3, which corresponds to a downsampling by a factor of three. We chose the scales to match our visual impression but they will need further validation.

III. RESULTS

To illustrate the methods described in Section II, we use the sample image [14] depicted in Fig. 1(a). We chose this picture because under phosphenated vision it can be challenging to see the different faces in front of a brighter background when relying on conventional methods for phosphenization. Our general goal is to facilitate the detection of faces (or another arbitrary feature) under phosphenated vision. Just by comparing the different phosphenization methods displayed in Fig. 1, the ideal accentuation method seems to depend on the number of phosphenes that are presented. For a low-acuity prosthesis, using elliptical phosphenes (Fig. 1(d)) can help in determining the number of faces. Conversely, a bright outline becomes more visible in a high-resolution image (Fig. 1(g)).

To validate this observation, we calculate the saliency map for each of the images. They are displayed in Fig. 2 and discussed in Section IV. Bright regions in the saliency plots mark the features that, according to this bottom-up model, are visually attractive. As this model does not take any training or prior knowledge into account, this might not display what an implantee would actually focus on and can only be a first approximation if it is generally easier to attend to the features we want to accentuate. In both phosphene resolutions, single phosphenes are salient. It also becomes clear that the algorithm prefers bright areas over dark areas. Even though contrast is the important feature, areas that are dimmer than a tenth of the maximal brightness are disregarded by the algorithm.

IV. DISCUSSION

As stated previously, bottom-up saliency can only help us pick up what appears to attract instantaneous attention. Thus, this algorithm is designed not to pick up faces specifically but to react to brightness, color and orientation. Looking at the low-resolution plot in Fig. 2(a), it becomes clear that faces do not naturally pop out of the image when we use the conventional methods to create phosphenes. What becomes salient according to the saliency model used are the bright background behind the faces and a bright phosphene in the bottom left that does not carry any important information. To increase the saliency of the faces, the outlines are

accentuated with maximally bright phosphenes. In a low-resolution image, this creates the problem that the three faces merge into one bright area. This can also be seen when looking at the computed saliency map in Fig. 2(b). The outlines become salient and it becomes evident that there is a region of interest. A different result can be seen in Fig. 2(c). Combining brightness with orientation, only the upper outline of the faces is salient. While this does not reflect what an implantee with prior knowledge about where to look would focus on, this result is still encouraging. That the accentuated feature pops out in a bottom-up way will make it easier for the user to concentrate on this image part.

We see different results for a high-resolution phosphene image. The saliency map for the conventional strategy (Fig. 2(d)) shows that the faces are not visually attractive and that, with a bottom-up model, the background would attract attention. However, a bright outline around the face is very salient and can thus direct the user to focus on these regions. Other than in the low-resolution case, the number of phosphenes is still sufficient to tell different outlines apart and a huge amount of grayscale information is preserved in the rest of the image. Figure 2(f) shows the saliency map of Fig. 1(h). It strongly resembles the saliency map for the conventional phosphenization method. Elliptical phosphenes at this scale are much harder to pick up, which makes them less salient.

We thus conclude that which accentuation technique is the most suitable depends on the number of phosphenes. While we see that oriented elliptical phosphenes might help to highlight features in low-resolution prostheses, outlining a feature with bright phosphenes is the more promising method of accentuation for a high-resolution prosthesis.

Looking at bottom-up saliency can only be a first step in evaluating these different methods. It does not reflect the human ability to focus on a specific target. Thus, we want to set up experiments with human observers for further validation of these preliminary results. The phosphene model that we use might not be a very accurate representation of what a patient actually perceives. It is lacking jitter in the phosphene positions and possible interactions between the stimulating electrodes. Also, it is not clear to what extent the brightness can be controlled and if we have to account for fading effects.

V. CONCLUSION

We presented two different methods for feature accentuation in phosphenated vision and used them for the example of faces. For each method, we found outlines of faces that we wanted to accentuate and displayed them with either maximally bright or elliptical oriented phosphenes. Our results show that which method is more promising depends on the number of electrodes and thus number of phosphenes. Because we lose some information by making the outline brighter, this method is less suited to low-resolution implants. However, elliptical phosphenes might not be easy to detect among a large number of round phosphenes, which makes using a bright outline the better method for a high-acuity implant. An application for high-acuity implants might lie in the combination of elliptical phosphenes with groundplane segmentation or simple edge detection, as it gives us the possibility of presenting edges and brightness information at the same time.

REFERENCES

[1] M. S. Humayun, J. D. Weiland, G. Y. Fujii, R. Greenberg, R. Williamson, J. Little, B. Mech, V. Cimmarusti, G. V. Boemel, G. Dagnelie, and E. de Juan Jr., "Visual perception in a blind subject with a chronic microelectronic retinal prosthesis," *Vision Research*, vol. 43, no. 24, pp. 2573 – 2581, 2003.

[2] S. H. Greenwald, A. Horsager, M. S. Humayun, R. J. Greenberg, M. J. McMahon, and I. Fine, "Brightness as a function of current amplitude in human retinal electrical stimulation," *Investigative Ophthalmology & Visual Science*, vol. 50, no. 11, pp. 5017–5025, 2009.

[3] A. Horsager, S. H. Greenwald, J. D. Weiland, M. S. Humayun, R. J. Greenberg, M. J. McMahon, G. M. Boynton, and I. Fine, "Predicting visual sensitivity in retinal prosthesis patients," *Investigative Ophthalmology & Visual Science*, vol. 50, no. 4, pp. 1483–1491, 2009.

[4] C. O. Savage, D. B. Grayden, H. Meffin, and A. N. Burkitt, "Predicting phosphene elicitation in patients with retinal implants: A mathematical study.," in *Proceedings of the 33rd Annual Engineering in Medicine and Biology Conference (IEEE-EMBS)*, pp. 6246 – 6249, August 2011.

[5] A. Horsager, R. J. Greenberg, and I. Fine, "Spatiotemporal interactions in retinal prosthesis subjects," *Investigative Ophthalmology & Visual Science*, vol. 51, no. 2, pp. 1223–1233, 2010.

[6] R. Wilke, G. Moghadam, N. Lovell, G. Suaning, and S. Dokos, "Electric crosstalk impairs spatial resolution of multi-electrode arrays in retinal implants," *Journal for Neural Engineering*, vol. 8, June 2011.

[7] C. O. Savage, F. I. Kiral-Kornek, and B. Tahayori, "Can electric cross-talk be used to control perception of a retinal prosthesis patient?." Submitted to 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'12), 2012.

[8] M. Hollins, *Understanding blindness: An integrative approach*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc., 1989.

[9] N. Tran, S. Skafidas, J. Yang, S. Bai, M. Fu, D. Ng, M. Halpern, and I. Mareels, "A prototype 64-electrode stimulator in 65nm CMOS process towards a high density epi-retinal prosthesis," in *Proceedings of the 33rd Annual International IEEE Engineering in Medicine and Biology Society Conference, (IEEE-EMBS)*, pp. 6729 – 32, 2011.

[10] E. Zrenner, K. U. Bartz-Schmidt, H. Benav, D. Besch, A. Bruckmann, V.-P. Gabel, F. Gekeler, U. Greppmaier, A. Harscher, S. Kibbel, J. Koch, A. Kusnyerik, T. Peters, K. Stingl, H. Sachs, A. Stett, P. Szurman, B. Wilhelm, and R. Wilke, "Subretinal electronic chips allow blind patients to read letters and combine them to words," *Proceedings of the Royal Society B: Biological Sciences*, 2010.

[11] D. Walther and C. Koch, "Modeling attention to salient proto-objects," *Neural Networks*, vol. 19, no. 9, pp. 1395 – 1407, 2006.

[12] J. Kim, X. He, and N. Barnes, "Automatic face zooming and its stability analysis on a phosphene display." Poster presentation at 2nd International Conference on Medical Bionics: Neural Interfaces for Damaged Nerves, November 2011.

[13] C. McCarthy, P. Lieby, and N. Barnes, "Ground surface segmentation for navigation with a low resolution visual prosthesis," in *Proceedings of the 33rd Annual International IEEE Engineering in Medicine and Biology Society Conference, (IEEE-EMBS)*, pp. 4457–60, 2011.

[14] H. Hayes, "Mount Rushmore." URL: http://fr.fotopedia.com/items/flickr-3734986331, 2009.

[15] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254 – 1259, 1998.

[16] L. Itti and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research*, vol. 40, pp. 1489 – 1506, 2000.

[17] J. K. Tsotsos, S. M. Culhane, W. Y. K. Wai, Y. Lai, N. Davis, and F. Nuflo, "Modeling visual attention via selective tuning," *Artificial Intelligence*, vol. 78, no. 1 - 2, pp. 507 – 545, 1995.

[18] F. I. Kiral-Kornek, C. O. Savage, and D. B. Grayden, "The focus of attention under phosphenated vision through retinal implants," in *Proceedings on Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, pp. 113 – 118, December 2011.